THE NATURAL PROBABILITY THEORY OF STEREOTYPES

– Jacob Stegenga –

Abstract: A stereotype is a belief or claim that a group of people has a particular feature. Stereotypes are expressed by sentences that have the form of generic statements, like "Canadians are nice." Recent work on generics lends new life to understanding generics as statements involving probabilities. I argue that generics (and thus sentences expressing stereotypes) can take one of several forms involving conditional probabilities, and these probabilities have what I call a naturalness requirement. This is the natural probability theory of stereotypes. Each of the two components of the theory entails a family of fallacies that contributes to the spurious reinforcement of stereotypes: inferential slippage within and between the different generic forms, and inferential slippage from facts about frequencies of group traits to beliefs about natural propensities or dispositions of groups. Empirical research suggests that we often commit these fallacies. Moreover, this theory can referee a vitriolic debate between some psychologists, who hold that stereotypes are always false and stereotyping is always wrong, and other psychologists, who hold that stereotypes are often accurate and stereotyping is often reasonable.
Keywords: stereotypes, generics, probability, naturalness, reasoning fallacies

1. Introduction

A stereotype is a belief or claim that a group of people has a particular feature. Stereotypes appear to be heuristics in human cognition. Conversely, stereotypes are often pernicious, and stereotyping contributes to various forms of oppression. Thus, understanding stereotypes can shed light both on cognitive processes and on social injustices. For these reasons, stereotypes have long been studied by social psychologists, and recently by philosophers. Here I offer a theory of stereotypes — the *natural probability* theory of stereotypes — and I articulate two families of fallacies associated with stereotypical reasoning which are predicted and explained by this theory of stereotypes.

Stereotypes are expressed by sentences that have the form and semantic properties of generic statements, such as "tigers are striped," "ducks lay eggs," and "mosquitos carry malaria." One approach to understanding stereotypes, then, is to understand generics. Generics have puzzled linguists and philosophers for a variety of reasons,

Jacob Stegenga
University of Cambridge
Department of History and Philosophy of Science
jms303@cam.ac.uk

prominently because the proportion of objects in a group that has the relevant property in generic statements varies widely — consider the tiny proportion of mosquitos that carry malaria, compared with the middling proportion of ducks that lay eggs, or the high proportion of tigers that are striped, yet the generics "mosquitos carry malaria," "ducks lay eggs," and "tigers are striped" are all true. Analyzing generics in terms of probabilities has been routinely dismissed by philosophers, linguists, and psychologists who study generics. However, recent work on the semantics of generics renders a probabilistic approach viable, and recent work on the psychology of stereotypes suggests that such an approach could be insightful. The first ambition of this paper is to analyze generics as probabilities, which then serves as the foundation for the natural probability theory of stereotypes.

I argue that the logical form of generics is polysemous, and their various meanings can be understood by appeal to probabilities. Generics can be understood as claims that are represented by one of several probabilistic forms. More specifically, generics (and thus sentences expressing stereotypes) can take one of several forms involving conditional probabilities. Pernicious stereotypes are often false because they do not meet the truth conditions of one or all of the generic forms. Even when a stereotype is true on one of the probabilistic forms, it can be false on a stronger form. The form that a speaker intends can be different than the form that an interlocutor infers, giving stereotypes their pernicious slipperiness. These conditional probabilities should be understood as arising from relatively constraining facts about groups, hence the natural probability theory of stereotypes.

This analysis suggests two reasoning fallacies that can contribute to the spurious reinforcement of stereotypes. Given the well-known tendency for people to reason fallaciously with probabilities, one fallacy of stereotypical reasoning involves inferential slippage between the different generic forms. Another fallacy involves inferential slippage from contingent facts about groups to beliefs about dispositions of groups. The psychological study of reasoning about probabilities and kinds shows that these two fallacies are ubiquitous. The first fallacy entails that a feature that is not held by many members of the group can come to be widely believed to be more ubiquitous than it in fact is. The second fallacy entails that such features can come to be wrongly thought of as natural dispositions of the group. A few polite Canadians and we soon have "Canadians are nice," a few quiet and thoughtful professors and we soon have "professors are absent-minded." The second ambition of this paper is to articulate these two fallacies of stereotyping.

Many people, including some social psychologists who study stereotypes, hold stereotypes to be by definition false, and stereotyping to be clearly wrong. Since most stereotypes cannot be understood to be true about all members of the group being described — obviously, not all Canadians are nice — they must be understood as claims about some proportion of the group being described (a fact which motivates the idea that stereotypes are generics, and which further motivates my probabilistic analysis) — but still, a common view is that even these claims are nonetheless false. A large empirical literature in social psychology is aimed at discovering biases that arise from stereotypes — for example, teachers with prior views about the promise of her students, based on

stereotypes, can inadvertently deploy more of her resources to those students with pre-sumed promise. A view often expressed by lay people and scientists studying stereotypes is that we should never stereotype people, because that involves treating people not as individuals but as a group, and since we should treat individuals as individuals (goes this view), stereotyping is wrong.

Other people, including some social psychologists who study stereotypes, argue that many stereotypes are relatively accurate depictions of social facts, and that relying on such stereotypes when making inferences about individual people can render such inferences more reliable. Stereotypical inferences are just like many other inferences in which we rely on probabilistic information about categories, goes this view, and to neglect such information when it is available is unreasonable. Whether or not any given stereotype is true or false is merely an empirical question, on this view, and we ought not pre-judge stereotypes as right or wrong on moral or other grounds. Instead, goes this view, we should judge stereotypes only by a clear-headed appeal to objec-tive facts regarding the extent to which people's beliefs about groups track relevant social facts.

These two positions on stereotypes contribute to a vitriolic debate which per-meates the scientific study of stereotypes. The natural probability theory of stereotypes offers an analysis of how stereotypes should be assessed, which can help to resolve this debate. That is the third ambition of this paper.

There are other ways in which stereotypes and stereotyping can be wrong — ethical, practical, or political wrongs — yet my hope is that by articulating in detail these two classes of epistemic wrongs of stereotypes some insight is gained on this consequential topic.

2. Stereotypes Are Generics

Not just any sort of belief or claim about groups is a stereotype. Stereotypes have a few features that distinguish them from mere descriptions of groups. "Canadians elect a prime minister" states that a group of people has a particular property, though the claim is not a stereotype — stereotypes tend to articulate presumed behavioral or dispositional properties of members of social groups, rather than properties that are, say, enshrined by laws governing those groups (as is the case with the Canadian electoral system). Often, the properties being attributed to members of a group in a stereotype are negative or dis-valued by the holder of a stereotype or the interlocutors sharing a stereotype. This need not always be the case — think of "Canadians are nice" — though it is a feature that makes many stereotypes pernicious. The expression of stereotypes can contribute to oppressive social practices.[1]

Some psychologists and philosophers have argued persuasively that stereotypes can be expressed as statements that have the form and content of generics: both generics and sentences expressing stereotypes are generalizations about kinds, both omit explicit

---

[1] Some have argued that this is the case even for positive stereotypes; see, for example, McKeever and Sterken (2021), Chestnut and Markman (2018), Foster-Hanson et al. (2016).

quantifiers such as "some" or "all," and both admit counter-instances.[2] More specifically, sentences expressing stereotypes are a subset of generics: sentences expressing stereotypes are generics about human groups, and these groups are typically demarcated by national, ethnic, and demographic properties (among others).

Stated this way, the notion of stereotypes is very broad — it includes many generalizations about groups that some might think are not stereotypes, such as "humans have hair." To restrict the notion further, some hold that stereotypes are cognitive heuristics that can be useful for explanation and prediction. So, for example, "Argentinians have mass" is not a stereotype, despite the fact that it is a generic about a human group, because it does not ground many useful predictions or explanations. To restrict the notion further still, some hold that stereotypes are generalizations about groups that have evaluative content (typically negative). However one decides to restrict the notion, stereotypes can be asserted as generics. So, in short, one way to understand stereotypes is to understand generics. That is the ambition of this section and the next.

The most salient feature of generics for this paper is that they can be represented as statements about probabilities — this is a controversial claim about generics, and I defend it in the following sections. Here I note a few preliminaries about generics that are worth articulating en route to a theory of stereotypes.

Cohen argues that generics are claims about groups which are akin to laws of nature.[3] Consider "tigers have stripes." What makes this statement true or false is a deep fact about biology: the constraints on felid fur are a result of millions of years of evolution. The same is true about other motivating examples, like "dogs have four legs," "mosquitos carry malaria," or "birds fly." Law-likeness is a matter of degree. The examples thus far have been about biological groups, and thus are governed by relatively fundamental laws, but those laws pertain only to particular groups of animals on Earth (as far as we know). Consider two other extremes of law-likeness. "Celestial objects within a planetary system orbit the barycentre of the system via an elliptical orbit" — this is a statement about all celestial objects at all times in all places, and thus it is an exemplary law of nature. "Smart phones are designed to be obsolete within a couple of years" — this is a generic statement about a very small class of objects, the existence of which was highly contingent on a number of cosmic accidents.

Despite affirming their law-like status, Cohen argues that generics are contingent rather than necessary — though a generic may be true in this world, they may be false in other worlds. We can conceive of other worlds in which tigers do not have stripes. Or consider the statement "spices are affordable." This is true only in a spatiotemporally restricted sense: spices are affordable to middle-class westerners today, but were not so in most historical periods.[4]

---

[2] This view is indeed widespread. See, for example, Hammond and Cimpian (2017), Beeghly (2015), Haslanger (2012); see also Begby (2013), Silva (2020), and Bosse (2022). Of course, stereotypes can be expressed in non-propositional forms, such as the depiction of particular images. Moreover, some stereotypes are about a putative group ideal ("women are empathetic"); this paper is concerned with stereotypes that are about a putative matter of fact ("Canadians are nice").

[3] Cohen (1999).

[4] Arguably, all physical laws are metaphysically contingent, though this is a matter of debate: necessitarians hold that physical laws are metaphysically necessary.

A related feature of generics is that they are often claims about groups that seem relatively natural. The groups being described in our examples above are dogs, mosquitos, celestial objects, and tigers, and not arbitrary collections of objects, such as all the objects on my desk today. However, think of "domestic pets are cute": the category "domestic pets" is hardly a grouping that tracks significant theoretical divisions in nature, and thus is not a natural kind; moreover, generics can be about artefacts, as in the example of smart phones above. Similarly, the property being predicated in a generic is often natural: emeralds are green, not grue — though again, the smart phone example suggests that this is not a strict feature of generics.

So generics involve groups that are often but not always natural, and properties that are often but not always natural. A more fundamental feature of generics involves various constraints causing members of the group to have the property. Haslanger argues, for example, that generic statements typically presuppose or imply essentialism regarding the ascription of the property to the group: Gs have F in virtue of the nature of Gs, and F is an essential property of Gs, goes this thought.[5] However, while Haslanger is right that many people suppose that generics imply essentialism (and there is empirical evidence that supports this, discussed below), the naturalness of generics (and thus stereotypes) cannot be the requirement of essentialism, as some of our examples already suggest: not all dogs have four legs, and smart phones could be designed to last longer than they do. The truth or falsity of a generic is undergirded by a variety of facts that are more or less constraining, but need not be as constraining as essentialism requires. I will call this aspect of generics *naturalness*. The kinds of facts and their corresponding degree of constraint that are sufficient such that one is willing to assent to a generic plausibly varies depending on context. The generic "humans have 23 chromosomes" is undergirded by deep biological facts operating at an evolutionary timescale, while the generic "spices are affordable" is undergirded by contingent social facts. I say more about the naturalness requirement below and in §4.

A related feature of generics is that it is neither sufficient nor necessary that a majority of a group (G) have a feature (F) for the generic "Gs are F" to be true. Cohen gives several examples: "Israelis live on the coastal plain," "people have black hair," and "books are paperbacks" are all claims in which the majority of the group does in fact have the stated property, though the claims sound false. Conversely, "dogs give birth to live young" is true despite the fact that fewer than half of all dogs give birth to live young. The generic "mosquitos carry malaria" is even more striking in this regard.

For the above reasons, psychologists and philosophers distinguish between generic statements and mere statistical statements. This is why, for example, you are probably willing to accept the statistical claim that "the majority of books are paperbacks" but not the generic claim that "books are paperbacks." The former is true but the latter is false, precisely because there is more to a generic claim than consistency with mere statistical facts. It is not in the nature of books to be paperbacks, even though most of them are paperbacks; conversely, it *is* in the nature of mosquitos to carry malaria, even though most of them do not carry malaria. Some have taken these considerations to imply that

---

[5] Haslanger (2012).

generics cannot be understood probabilistically, but this follows only on a narrow understanding of probability (namely, an actual frequency interpretation of probability), as I argue below, and on a dubious assumption about generics.

In Sterken's insightful account of generics, she notes that a commitment of many theorists is an "assumption of unity" regarding generics, which holds that "there is a unified phenomenon of genericity that generic sentences, in general, instantiate."[6] Against this, Sterken argues that we should accept that generics have truth-conditional variability. Sterken makes a convincing case that generics are context-sensitive indexicals. Just as the meaning of words like "you," "she," "here," and "now" depend on context, the meaning of generic statements "Gs are F" depends on context. This can be called "contextualism" about generics.

Two features of generics can vary by context on Sterken's account: their domain-restriction and their quantificational force. What Sterken means by quantificational force seems to be the proportion of members of a group which have the property expressed in the generic. This notion of quantificational force suggests a straightforward representation by probabilities. Sterken argues that the semantic value of a generic, and especially its quantificational force, is determined by the intentions of the speaker of a generic and what a competent listener who understands the common ground of the conversation would infer about the speaker's intentions. This will be important in the next section, in which I argue that generics can take multiple probabilistic forms, and so for any given generic one must determine which form applies.

As suggested above, another feature of generics that can vary by context is the degree of worldly constraints on the group such that members of the group in fact have the ascribed property. I have been calling such constraints "naturalness." This term may be misleading, because it might suggest that the constraining facts are always natural facts — say, facts about the evolutionary history of an animal species. While this may be the case in the examples above about animal species, the constraining facts relevant to generics about, say, technological artefacts or human social groups may be less fundamental. So the naturalness requirement varies in degree, and what determines the strength of the requirement for any given generic is, as on Sterken's analysis, the intentions of the speaker of a generic and a competent interlocutor's assessment of those intentions, based on common ground.

3. Generics and Probabilities

Generics share a superficial form in natural language: "Gs are F." The semantic content of generics, however, varies widely. Generics can express multiple kinds of probabilistic generalizations.[7] These generalizations can be represented with conditional probabilities. Conditional probabilities are of the form "the probability of A, given B" — the probability that it is raining outside, given that I am in England, for example. To express such statements efficiently, one writes $P(A \mid B)$ to represent the probability of A given B.

---

[6] Sterken (2015a). See also Nickel (2012) for suspicion that there is a "single semantic core" underlying the diversity of generics.

[7] Many generics can be represented by other kinds of statements, such as quantified statements, like "many Canadians are nice," but this is of course compatible with a probabilistic approach.

The first account of generics as conditional probabilities was offered, as far as I know, by Cohen. In this section I extend Cohen's account of generics, and in the next section I defend it against several objections.[8]

I will use a simple notational convention: a generic statement that claims that group (G) has feature (F) will be expressed with the conditional probability $P(F|G)$: the probability that some entity has feature F given that the entity is in group G. Conversely, the term $P(G|F)$ expresses an "inverse probability": the probability that some entity is in group G given that the entity has feature F. Those two conditional probabilities are not the same. $P(F|G)$ should not be equated with $P(G|F)$ — to do so is to commit a fallacy. Later this will be important.

The form itself of generic statements is polysemous. Consider the following examples: "dogs are mammals," "dogs have four legs," "dogs give birth to live young," "dogs have canine transmissible cancer," and "dogs are violent." These all share the superficial generic form, in which a particular group is described as having a property, but their meanings vary widely. I will use D for the group (dogs) and M, L, Y, C, and V for the five ascribed properties, respectively (mammalian, four-legged, giving birth to live young, having canine transmissible cancer, and being violent). For each generic form I state its representation in terms of conditional probabilities.

Universal Generic
"Dogs are mammals"
$P(M|D) = 1$
The idea: these are universally quantified statements. All dogs are mammals, as a result of deep biological constraints.

Absolute Generic
"Dogs have four legs"
$P(L|D) > x > 0.5$
The idea: statements like this claim that the majority of a group has a particular property. Most, but not all, dogs have four legs. The extent of the majority that is required for an absolute generic to be true is, at a minimum, 0.5, though in many contexts x will be much higher.

Relative Generic
"Dogs give birth to live young"
$P(Y|D) / P(Y|D^*) > y > 1$
The idea: Since it is not the case that the majority of dogs give birth to live young, generics like this cannot be understood as absolute generics. Still, it is true that dogs give birth to live young. For such generics, a higher proportion of the group has the feature compared to a contrast group, D*.

---

[8] Cohen (1999). Cohen's account includes only absolute generics and relative generics. For the reasons I give below, to understand stereotypes and to adequately respond to alleged counter-examples to a probabilistic account of generics, I include three other kinds of generics.

Unique Capacity Generic
"Dogs have canine transmissible cancer"
$P(D\,|\,C) = 1$
The idea: The claim expresses the fact that only dogs have canine transmissible cancer. Notice that the conditional probability for unique capacity generics is the inverse of the others — it is based on $P(G\,|\,F)$, rather than on $P(F\,|\,G)$. (I am unsure if many stereotypes are unique capacity generics, but because, as we will see below, such generics have been cited as challenges to a probabilistic analysis of generics, I include them here.)

Capacity Generic
"Dogs are violent"
$P(V\,|\,D) > 0$
The idea: the probability that some member of the group has the property must be greater than zero. This condition is very weak. Nevertheless, having this form will help the discussion of stereotypes, because many stereotypes will turn out to be true only in this most trivial sense.

If a generic can be represented by one of the above forms, how can we know, for any asserted generic, which form applies? Sterken offers us an answer.[9] Recall that, on Sterken's contextualism about generics, the meaning of a generic, and especially what Sterken calls the "quantificational force" of a generic, is determined by the intentions of the speaker of the generic and what a competent interlocutor understands about the speaker's intentions. So, the generic "Gs are F" can be a universal generic, an absolute generic, a relative generic, a unique capacity generic, or a capacity generic, and to tell which it is involves making an informed inference about the speaker's intentions by appealing to context and common ground between speaker and interlocutor.[10]

Informed and attentive interlocutors ought to (and often do) ascribe to the speaker of a generic the form of the generic that the speaker intends, and that is often, but not always, the form that renders the generic true. When Valentyna claims "mosquitos carry malaria," I ascribe to her the assertion that some mosquitos carry malaria, or that only mosquitos carry malaria, and not that most mosquitos carry malaria. How do I know to ascribe this intended assertion to Valentyna? Our shared common ground and context. Valentyna is educated, honest, and knows much about tropical diseases, and so I interpret her assertion as a generic of the form which renders it true. When Bob, a liberal sociologist who studies the frequency of violent crimes in different racial groups, claims "blacks are violent," I take him to be asserting a relative and not absolute generic, and I take him to be making a contingent claim rather than a claim about the group's deep-seated dispositions (more on this below); when Serge, a member of a racist organization, makes the same claim, I might take him to be asserting an absolute generic, and

---

[9] Sterken (2015a).
[10] Empirical studies show that people understand that generics represent a wide range of prevalence levels, and this has been shown in subjects as young as pre-school children (Brandone et al. 2015). Nguyen (2020) has recently defended a view of generics along similar lines, in which the speaker's intentions determine the kind of quantified expression being asserted in a generic.

I might take him to be asserting a claim about the dispositions of members of the group (and thus I would be ascribing to Serge the assertion of a generic which is false, though Serge may wrongly believe it to be true).

This list of generic forms is not arbitrary. When Antara says "dogs have four legs," she means to assert (and an attentive interlocutor would take her to mean) "most dogs have four legs," and thus she means to assert an absolute generic. When Sveta says "ducks lay eggs," she means to assert (and an attentive interlocutor would take her to mean) that the proportion of ducks that lay eggs is larger than the proportion of individuals in a relevant and well-chosen contrast group. These five representations of generics cover all examples of bare plural generics that I am aware of (though as noted earlier, I am putting aside generics that appear to assert ideals rather than factual descriptions, such as "women are empathetic"). Moreover, quotidian linguistic practice and psychological research both suggest that the proportion of a group which has a feature expressed in a generic which is accepted as true can vary widely.[11]

My analysis thus far has relied on stipulations that when a person P asserts a generic G, they mean Z. Yet, one might object that, for all I know, when P asserts G they mean Z*. My stipulations that P means Z rather than Z* have been based on considerations such as the fact that G could only be true if understood as expressing Z rather than Z*, since the facts of the world warrant Z and not Z*. Nevertheless, we do not have special access to P's intentions, beyond all the usual features of interpersonal communication that allow an interlocutor to make an inference about a speaker's intended meaning. As we will see below, the possibility that P means Z* rather than Z is precisely one of the features of stereotypes that contributes to the too-frequent adoption of false stereotypical beliefs and the widespread misunderstandings associated with assertions of stereotypes. Nevertheless, as argued, shared context can often allow an interlocutor to understand that a speaker of a generic means Z when they indeed mean Z.

This contextual approach to understanding generics is relevant to other aspects of my analysis. For instance, the value of x in an absolute generic can be understood contextually. "Dogs have four legs" assumes or implies a very large value of x, whereas "dogs are cute" assumes or implies a smaller value of x. Similarly, the specific contrast group in a relative generic can be understood contextually, as can the value of y.

Stereotypes can be contrasted with prototypes. Prototype theory was introduced by Rosch in her study of the cognition of categories.[12] A prototype is the "most typical" member of a category. Thus, the prototype dog would have four legs. So prototypes only take an absolute generic form, whereas, as we have seen, stereotypes can take a variety of forms.

Probabilities have several possible interpretations. One interpretation holds that probabilities are objective frequencies, another interpretation holds that probabilities are subjective beliefs, while a third holds that probabilities are propensities. Some theorists maintain that all statements of probabilities must be interpreted according to only one of these interpretations, while others maintain that different probability statements can be understood by different interpretations. What kind of probabilities are generics?

---

[11] See, for example, Cimpian et al. (2010), Tessler and Goodman (2019).
[12] Rosch (1973), especially p. 330.

As we saw in §2, some generics are statements regarding relatively deep facts about groups: the features ascribed to groups in the biological examples tend to be relatively natural, the features tend to be relatively intrinsic properties of groups, the groups tend to be relatively natural collections of individuals, and the natural properties of the groups tend to be causally relevant to the ascribed features. Generics tend not to be based merely on contingent facts, like "books are paperbacks," spurious properties, like "trees are grue," or contingent groupings, like "Canadians are right-handed." We saw earlier that the probabilities representing generics cannot be mere statistical summaries of contingently occurring facts, which rules out a naive frequency interpretation of probabilities for generics. For such reasons, Cohen prefers a long-run hypothetical frequency interpretation. Another option is to understand such probabilities as propensities, and indeed, the propensity interpretation of probability is sometimes characterized as long-run hypothetical frequency.[13] Still another option would be to understand such probabilities epistemically, in which the probability statements would encode missing information; for example, "ducks lay eggs" may be a relative generic, but "healthy and mature female ducks lay eggs" could be understood as an absolute generic with a high value of x.

4. A Theory of Stereotypes

Everything is in place, now, for a theory of stereotypes. Since statements expressing stereotypes are a subset of generics, and generics are statements making claims about probabilities, stereotypes are statements making claims about probabilities. Moreover, stereotypes have a probability condition and a naturalness condition. More specifically, stereotypes are expressed as statements which take one of the above forms involving conditional probabilities, and these statements hold (or do not hold) in virtue of the extent to which facts about the group in question constrain the ascribed property. Thus there are two necessary conditions for a stereotype to be accepted as true: the facts about the group in question must warrant one of the probabilistic generic forms articulated above, and those facts must ground such warrant as a result of dispositional or other constraining facts about members of the group. Conversely, there are two ways in which a stereotype can be false: if it does not satisfy the probability requirement or if it does not satisfy the naturalness requirement (I offer more detail on the naturalness condition below). As we will see, a great deal of ambiguity remains in this account of stereotypes, but this ambiguity is faithful to the way people in fact reason with stereotypes, and indeed, as I argue below, it is this ambiguity which affords the fallacious reasoning so often exemplified with stereotypes.

Many stereotypes will turn out to be false under any of the generic forms. Some stereotypes may in fact seem to be true absolute generics — examples might include "Canadians are nice," "Danes are liberal," and "Brits have bad teeth" — but many such examples are true in the merely statistical sense noted in §2, and not in the generic sense,

---

[13] The propensity interpretation of probability has faced a number of challenges, perhaps most notably that of Humphreys (1985), but these challenges have been recently addressed by Berkovitz (2015). See also Suarez (2020).

and thus ought to be rejected. Some stereotypes may turn out to be true on one form (say, as a relative generic) but false on another (say, as an absolute generic) — think of "Asians are good at math" or "Russians are good ballet dancers" (if there is any contingent factual basis to such stereotypes, they are relative generics which probably involve small *y* values).[14]

Despite ambiguity about the intended generic form and intended values of *x* and *y*, at the very least the probability requirement articulated above is formulated with precision. The same cannot be said of the naturalness requirement. Consider some examples to illustrate the requirement. "Tigers have stripes" is an absolute generic, and it is true in virtue of deep biological facts about tigers. "Dogs give birth to live young" is a relative generic, and it is also true in virtue of deep biological facts about dogs. But what about "Danes are liberal"? How constraining do the relevant worldly facts have to be such that the naturalness requirement is satisfied, for such stereotypes? Unlike the probability requirement, there is no well-defined metric of naturalness, though more can be said about what the requirement demands.

Slater offers an interesting account of natural kinds, in which naturalness is articulated as a graded property and formulated using the tools of probability.[15] Slater remains quiet on the physical properties that ground naturalness; the important feature to consider is the extent to which properties cluster. Some property clusters are extremely spatiotemporally contingent — "the coins in my pocket are Ukrainian hryvnas" involves a fragile clustering of properties if I am on a short holiday in Ukraine. Other property clusters are more stable — "tigers have stripes" is a result of spatiotemporally durable biological facts.

The illustrative examples of the naturalness requirement have been about animal species, in which the requirement is satisfied as a result of constraints by facts on an evolutionary timescale. In general, however, one might hold that we should not demand so much of the naturalness requirement for stereotypes. If we did, we would rule out the vast majority of stereotypes (perhaps all) with one broad stroke, since stereotypes are claims about groups and about properties that are neophytes relative to the groups and properties involved in the illustrative generics about animal species. Stereotypes are about social kinds, not natural kinds. It would be a crude theory of stereotypes — goes this thought — if it ruled out "Russians are good ballet dancers," "Danes are liberal," and "Finns drink a lot of alcohol" by bluntly objecting that ballet, liberalism, and alcohol consumption are contingent and relatively new kinds of things (in the long run of the

---

[14] Indeed, in my analysis of relative generics I required only that *y* > 1, which can hold when the group has the feature at only a very slightly higher frequency than a relevant contrast class. For many stereotypes the assumed or implied value of *y* in a stereotype is in fact much larger than 1. Suppose we accurately assessed the proportion of redheads in every country, compared with all other countries, and for the three countries with the highest *y*-values, the *y*-values were:

Country 1   1.0007
Country 2   1.0006
Country 3   1.0005

I doubt that you would be willing to accept the stereotype "people from Country 1 are redheads." This suggests that for many stereotypes, a value of *y* significantly larger than 1 is assumed or implied.
[15] Slater (2015).

world), or that the existence of national groups like Russians, Danes, and Finns is recent and contingent. The naturalness requirement should not just rule out *all* stereotypes, goes this consideration.

On the other hand, one might think that the naturalness requirement must rule out *some* stereotypes, otherwise it is no requirement at all. I suggest merely that, like probabilities, the naturalness requirement comes in degrees, though unlike probabilities, we cannot say much about its scale. Anyway, to illustrate, consider:

(1)   "Russians vote for a person whose first name is Vladimir"
(2)   "Russians are good ballet dancers"
(3)   "Russians have pale skin"

These claims are grounded on progressively less contingent and progressively more constraining biological, historical, and sociological facts.[16] For precisely this reason, I anticipate that many readers will hold that (1) is not a stereotype. That intuition can be explained by the natural probability theory of stereotypes: (1) satisfies the probability requirement but not the naturalness requirement. (1) is statistically true but not because there is a deep-seated disposition of Russians to vote for a person whose first name is Vladimir. On the other hand, I anticipate that readers will hold that (2) and (3) are stereotypes. Again, this can be explained by the natural probability theory of stereotypes: (2) obviously must be understood as a relative generic, and moreover, unlike (1), (2) is grounded by a cultural milieu that goes back centuries. (3) is grounded by even more constraining biological facts.

So the naturalness of stereotypes is a gradable property. The degree of naturalness is determined by the various kinds of constraints that ground the fact that the group in question has the ascribed property. I believe views will differ regarding what sorts of constraints are admissible or necessary when determining whether the naturalness condition is satisfied. To see this, consider Haslanger's example:

(4)   "Blacks are violent"

Suppose, for the sake of argument, that the data that suggests that blacks commit more violent acts than other racial groups in the United States is not merely a result of methodological biases of measurement (that this data *is* a result of such biases is of course a live possibility). Suppose further that to the extent that (4) is true, it is a result of centuries of systematic and severe oppression of blacks. Thus, if the naturalness requirement can be satisfied by cultural forces operating on a centennial timescale, (4) satisfies a reasonable interpretation of the naturalness requirement. And since, as supposed, (4) satisfies the relative generic condition, (4) would come out as a true stereotype. There are several responses that one might find persuasive. One might accept (4) as true while being explicit about the causes of its truth. Or one might reject (4) as false because the relevant naturalness requirement is too loose. On this latter response, one might say: cultural forces operating on a centennial timescale can only cause groups to have properties very contingently; the naturalness requirement should be stronger (perhaps by requiring deeper biological causes of traits, as is the case with many of the illustrative examples of generics from biological species). On this latter response, (2) would also come out false.

---

[16] So I claim, at least. One might disagree with me, which, as we will see below, would illustrate one of the families of fallacies of stereotypical reasoning.

The strengthening of the naturalness requirement, which rules out both (2) and (4), might strike some as too blunt, since arguably most social stereotypes do not satisfy a strong naturalness requirement. On the other hand, this approach might strike others as exactly correct, insofar as it shows that, to the extent that some groups do in fact have properties that are described in social stereotypes, they do so very often only contingently.

Consider this example:

(5)   "Californians are liberal"

This is a stereotype, but is it a *true* stereotype? (5) is true only if one holds a relatively loose naturalness requirement. In the last few presidential elections Californians have voted for the plausibly more-liberal candidates, and so the stereotype has some recent statistical warrant. Yet this warrant is fragile: Ronald Reagan won 53% of the California vote in 1980 and more than 57% of the vote in 1984, George Bush Sr won California in 1998, Bill Clinton won only 46% of the California vote against plausibly less-liberal candidates in 1992, and California's elected governors often hold illiberal political platforms. The naturalness condition directs attention to the degree of contingency of putative statistical truths, allowing us to challenge stereotypes like (4) and (5) that have factual warrant which is historically contingent.

A related consideration asks how we ought interpret the predicates used in stereotypes, such as "violent" in (4). With such behavioral predicates we can distinguish between actual behavior and behavioral dispositions. Consider Andrei, who has been a soldier fighting in eastern Ukraine for the last five years. He has been firing guns at others for most of his young adult life. But now he has returned home and he intends to live (and is predisposed to live) a peaceful, quiet life. The claim "Andrei is violent" is true only if the predicate is interpreted in terms of his actual recent behavior; the claim is false if the predicate is interpreted in dispositional terms. To know that the latter interpretation is false requires knowing something relatively deep about Andrei, namely, his behavioral dispositions. In short, we can interpret the claim as being either "Andrei has committed many violent acts in recent years" (which is true) or "Andrei is predisposed to be violent" (which is false). Complicating matters is the possibility that Andrei's recent violent context might have in fact influenced his behavioral dispositions.

Still another nuance about the strength of the naturalness condition is the point noted in §2: the presumed degree of constraint which grounds a stereotype can vary by context. It is the common ground between speaker and interlocutor which permits a shared understanding of the degree of constraint in place for any given stereotype. And in turn, the absence of sufficient common ground can afford misunderstanding regarding the degree of constraint.

Despite the imprecision of this notion of naturalness, it can be deployed in two ways. It can be used to criticize asserted stereotypes for which the naturalness requirement very clearly is not satisfied. Second, insofar as there are disagreements or misunderstandings about whether the naturalness requirement is satisfied, the mere articulation of the requirement can explain why stereotypes lend themselves to disagreements or misunderstandings (I develop these ideas in §6).

5. Objections Thus Far

There are several developed accounts of generics in the philosophical literature, the survey of which would take me astray.[17] In this literature a number of objections have been raised against a probabilistic approach to generics. These objections tend to be based on alleged counter-examples: examples that are intuitively true generics but come out as false on a probabilistic analysis, or examples that are intuitively false but come out as true on a probabilistic analysis. A fully developed probabilistic account of generics can diffuse these objections.

"Humans are autistic"
Here is an alleged counter-example from Leslie.[18] Humans suffer from autism, unlike other animals, though at a low frequency yet certainly at a higher frequency than other animals. The generic "humans are autistic" clearly cannot be understood as an absolute generic. However, it appears to satisfy the truth conditions of a relative generic, yet it seems intuitively false. Thus, goes this objection, there is a problem: the claim satisfies the truth conditions of a relative generic, yet the claim is untrue. Such an example can only be an objection if it is properly understood as a relative generic. However, I do not think "humans are autistic" ought to be understood as a relative generic. You might expect that an informed interlocutor who shares common knowledge with the speaker of "humans are autistic" will understand that the speaker intends to assert either "some humans are autistic" or "only humans are autistic," in which case the statement is best understood either as a capacity generic or a unique capacity generic. Its tenor is similar to "humans are violinists": true capacity generics and true unique capacity generics. The low but non-zero prevalence of autism among humans warrants the claim that some humans are autistic or that autistics are human, while the categorical impossibility of the property in any other contrast class renders the comparison of the proportion of autism among humans to other species meaningless.

Consider an extreme example. It is a requirement of the United States constitution that American presidents must be natural born citizens. The proportion of American natural born citizens who have been American presidents is of course very low, but it is higher than all other contrast classes, because in all other contrast classes it is zero. Now consider the claim which is analogous to the autism example: "American natural born citizens are American presidents." This is of course absurd. Based on our background knowledge, we know not to infer from the higher proportion of American presidents among American natural born citizens compared to other groups the relative generic "American natural born citizens are American presidents"; instead, we know to infer the unique capacity of American natural born citizens to be presidents: "American presidents are American natural born citizens." When the context entails that a generic claim is asserting a unique capacity of a group, the conditional probability is the inverse of that in a relative generic (§3).

---

[17] See, for example, Pelletier and Asher (1997), Leslie (2008), Asher and Pelletier (2012), Nickel (2016), and Sterken (2017).
[18] Leslie (2008).

There is, in addition, a convenient mathematical feature in the analysis of relative generics that blocks alleged counter-examples like the autism case. Recall that relative generics are represented by the ratio P(Y|D) / P(Y|D*). If one insisted on interpreting "humans are autistic" as a relative generic, one would be faced with the uncomfortable fact that, for any contrast class, the denominator of this ratio is zero, and thus the ratio would be undefined. Interpreting "humans are autistic" as a unique capacity generic does not face this problem.

To sum, the only way such an example is a problem for the probabilistic account of generics is if the claim is understood as a relative generic, because, for any chosen contrast class, the assertion satisfies the truth conditions of relative generics. But such claims simply should not be understood as relative generics. When a generic involves a claim that a group and only that group has a property, it is a unique capacity generic.

"Fleas carry malaria"

In Cohen's original analysis of generics, he understood generics such as "mosquitos carry malaria" as relative generics. On the above analysis, "mosquitos (M) carry malaria (R)" is a unique capacity generic: P(M|R) = 1. Cohen's approach works too: P(R|M) > P(R|M*), where M* is some salient contrast class (say, all other insects). Leslie proposes another counterexample to an interpretation of such claims as relative generics. Suppose fleas also carry malaria, but at a higher frequency than mosquitos, and the flea population increases such that fleas outnumber all other insects. Now the statement does not satisfy the relative generic probability condition (at least for one contrast class). But, goes this objection, it is compelling to think that the statement remains true — after all, notes Leslie, one can still catch malaria from mosquitos. We can no longer understand the statement as a unique capacity generic, because in the toy example, carrying malaria is no longer a unique capacity of mosquitos, and the inverse probability is quite small: P(M|R) < 0.5.

The importance of this case hinges on several substantive points: after the increase in the flea population whether the statement in fact remains true, what generic form an informed interlocutor would ascribe to the statement, and if it is interpreted as a relative generic, what the relevant contrast class is. Take the latter point first. On virtually all contrast classes (with only one exception) the statement will satisfy the relative generic condition. Consider a similar case: "university students are youthful." Sounds true to me, despite the fact that universities admit the occasional mature student — and it is true despite the fact there are even more youthful contrast classes, like kindergarten students. Alternatively, one might understand the speaker of "mosquitos carry malaria," after the increase in the malaria-carrying flea population, to be asserting a capacity generic. And indeed, when Leslie urges us to maintain that the statement remains true after the increase in the flea population, the only argument that she offers is that one can still catch malaria from mosquitos, which is to appeal to a capacity. After the flea population has increased, "mosquitos carry malaria" is like "humans climb rock walls": the latter is true of a very small proportion of humans, though a very large proportion of geckos, spiders, and mountain goats climb rock walls, and thus "humans climb rock walls" can be understood as a capacity generic and to understand it as a relative generic

would require a careful choice of contrast class. In any case, this thought experiment is not, in the end, an objection to a probabilistic approach to generics.

To make sense of generics like "mosquitos carry malaria," Leslie introduces the notion of striking property generics. These are generics which assert that a group has a salient, often dangerous, feature, even if the feature occurs at a low frequency in the group. But we do not need this notion to understand such cases. When we say "mosquitos carry malaria," what we mean is "only mosquitos carry malaria." Leslie's striking property generics creates trouble because of stereotypes like "Muslims are terrorists." Though the property is at a low prevalence for the group, being a terrorist is a striking property, and so her account faces the uncomfortable conclusion that the stereotype may turn out to be, on her account, true. To avoid this conclusion, in later work Leslie suggests that the problem with "Muslims are terrorists" is that the group has the property at the prevalence that it does only contingently rather than essentially.[19] That is, of course, a compelling position. But I have offered a more straightforward way to deny the truth of this stereotype. An informed interlocutor ought to understand that "Muslims are terrorists" is clearly false as a unique capacity generic (as opposed to "mosquitos carry malaria," which is a true unique capacity generic). Moreover, an informed interlocutor with a broad enough evidence base ought to understand that "Muslims are terrorists" is very likely false as a relative generic, too, and that it is obviously false as an absolute generic. So what kind of generic could it be? It is of course true when interpreted as a capacity generic, but that is a weak statement indeed, and almost certainly not what a speaker of such a pernicious stereotype intends to assert.[20]

This is not to say that Leslie's appeal to the contingency of some stereotypes is unimportant — as I argued above, the natural probability theory of stereotypes itself has a naturalness requirement. On the natural probability theory, however, stereotypes like "Muslims are terrorists" are false twice over, based on both the probability and the naturalness considerations, whereas on Leslie's account of generics such stereotypes come out as true but contingently so.

"Dutch people are good sailors"
Here is an objection to Cohen's probabilistic approach to relative generics, raised by Nickel.[21] Suppose some Dutch sailors are among the best sailors in the world, and the proportion of Dutch sailors who are among the best sailors in the world is higher than in other countries. However, as one would expect in any country, the majority of Dutch people are terrible sailors. Moreover, Nickel argues that with gradable relative proper-

---

[19] Leslie writes "One possibility is that these "dangerous generics" involve the assumption that the members of the kind in question share a nature that ground the property in question (or at least the corresponding disposition or propensity), and so generalizations such as "Muslims are terrorists" again involve false essentialist beliefs about social groups" (2016). See also Leslie (2017).

[20] Sterken (2015b) raises several objections to a capacity reading of generics, and argues instead for an "error theory" about generics that have been given a capacity reading, such as "sharks attack bathers." That is, Sterken argues that our intuitions that the relevant generics are true are wrong. This has the interesting result that any stereotype that can only be understood as a capacity generic on my account, such as "Muslims are terrorists," would be, according to Sterken's error theory, false.

[21] Nickel (2016).

ties like *being a good sailor* populations can be polarized, such that the majority of Dutch people are much worse sailors than, say, the majority of French people. Thus, goes this objection, the relative generic "Dutch people are good sailors" is intuitively false, though if we set a conventional standard for "good sailor," it could come out true according to the relative generic truth condition, and this is taken to be a problem for a probabilistic analysis of relative generics. However, I do not share the intuition that renders such cases counter-examples. It is compelling to understand generics like "Argentinians are good tango dancers," "Bulgarians are good weightlifters," and "Kenyans are good distance runners" as relative generics, despite the fact that the majority of people in these groups do not have the predicated feature, and even if, in fact, most people in these groups have the opposite feature. This is precisely because these generics clearly are not absolute generics, and Kenyans *are* good distance runners, Bulgarians *are* good weightlifters, and Argentinians *are* good tango dancers — even if the majority of the groups in question don't have the ascribed features, such generics can be relied on for useful explanations and predictions.[22]

Cohen motivates his account of generics by noting that absolute generics can be used to make useful predictions. Can relative generics provide such pragmatic utility? If you randomly chose a few people from the streets of London, put the Dutch in one boat and the Belarusians in another boat, the generic "Dutch people are good sailors" will not help you make a reliable prediction about whether or not the Dutch or the Belarusians will win the boat race. However, many relative generics can be pragmatically useful, because they can form the basis of compelling explanations, reliable predictions, and effective decisions. Suppose Anastasia is a tango dancer from Kyiv, and she wants to spend her winter holiday in a place in which she can have many good dances with skilled tango dancers. She recalls the relative generic "Argentinians are good tango dancers", and so she decides to spend her holiday in Buenos Aires. The vast majority of Argentinians are not good tango dancers, but this does not matter for Anastasia's practical ends — she doesn't intend to dance with most of the country's populace, but only a very small and selected subset of the country's populace. The relative generic is enough for Anastasia to make a reliable prediction.[23]

"This coin normally lands heads"
Here is a putative challenge to a probabilistic account of absolute generics:
"consider a slightly biased coin that comes up heads 50.000000001 % of the time. The probabilistic account predicts that *this coin normally comes up heads* is a true generic. This can't be right according to our intuitions."[24] It would indeed seem strange to say that this coin normally lands heads.

But this is no problem for the probabilistic account of generics presented in §3. Recall that absolute generics were defined according to a particular threshold proportion

---

[22] See also van Rooij (2019).
[23] This consideration suggests a refinement of the probabilistic analysis, initiated by Cohen (1999), in which the group is divided in salient partitions such that each partition is homogeneous with respect to a salient property, and the probabilistic inequalities hold for some or all of the partitions.
[24] Asher and Pelletier (2012).

x of the elements of the group that have the property, and that the particular value of *x* is determined by context. Whatever considerations inform one's intuition about the strangeness of asserting "this coin normally lands heads" can be the basis of an approximate value for *x*, presumably significantly higher than 50%. I share the intuition that an *x* of 50% for cases like this is too low. If a coin were biased to land heads 95% of the time, it would not seem so strange to assert the generic "this coin normally lands heads."

6. Two Fallacies of Stereotypes

This section develops Haslanger's insight that "attention to the ambiguities and slippages between different linguistic forms is useful in explaining how ideas become entrenched and social practices seem natural and inevitable."[25] The natural probability theory of stereotypes suggests two distinct families of errors in stereotyping. The first is a family of fallacies involving ambiguity in and slippage between the probabilistic generic forms of stereotypes. The second is a family of fallacies involving unwarranted conclusions about the naturalness of the ascribed feature, including unwarranted inferences from facts about frequencies to beliefs or claims about dispositions or propensities.

Formal slippage fallacies
We saw that shared background context between speaker and interlocutor helps one infer which generic form is being expressed by an assertion of a stereotype. Similarly, for absolute generics, background context helps one determine the approximate value of *x*, and for relative generics what the salient contrast class is, and the approximate value of *y*. This reliance on a complex tapestry of context and intentions affords ambiguity and slippage. A speaker might believe and intend to express a stereotype of one generic form, while an interlocutor might misunderstand the speaker as believing or expressing the stereotype as taking a different generic form. Or, both speaker and interlocutor might share an understanding that a stereotype is intended to be interpreted as an absolute generic, but the speaker might have one value of *x* in mind while the interlocutor has a very different value in mind. Or, both speaker and interlocutor might share an understanding that a stereotype is intended to be interpreted as a relative generic, while the speaker has one contrast class in mind and the interlocutor has another. Also with relative generic stereotypes, a speaker might have one approximate value of *y* in mind while the interlocutor has a very different value of *y* in mind.

People do not reason well with probabilities. For example, a common reasoning fallacy is to confuse a conditional probability with its inverse. That is, when presented with reasons to think that $P(A \mid B)$ is the case, many people fallaciously infer $P(B \mid A)$. This is sometimes referred to as the inverse fallacy. Empirical work by psychologists such as Kahneman and Tversky shows that people are prone to commit such fallacies. The ubiquity of probabilistic reasoning fallacies, together with my probabilistic analysis of generics, suggests one reason why stereotypes can be widely believed to be true even if they are false: people slip from one probabilistic form of stereotypes to another.

---

[25] Haslanger (2012): 449.

As argued earlier, inferring the correct form of an asserted generic, and the value of x (if absolute generic) or the value of y (if relative generic) requires an informed interlocutor sharing common ground with the speaker. In real contexts, an interlocutor is not perfectly informed and might not share much common ground with a speaker. One's inference about the generic form, the value of x (if an absolute generic), and the contrast class and value of y (if a relative generic) are all fallible.

I have offered a theory about how to understand statements asserting stereotypes, but this theory also suggests how stereotypes might initially develop. Some stereotypes begin life as a collection of facts that warrants merely a capacity generic, or perhaps a relative generic with a modest value of y. The holder of this generic or his interlocutor then wrongly infers a stereotype of a stronger form, such as a unique capacity generic or a relative generic with a larger value of y. Sometimes the inference goes as far as an absolute generic.

There are some empirical findings that support this theory of stereotypes and the associated fallacy of slippage between generic forms. In one experiment, subjects interpreted novel generic statements about fictive categories as referring to a large proportion of the members of the group in question: subjects interpreted "lorches have purple feathers" as referring to most lorches. Conversely, when presented with evidence that only a small proportion of lorches have purple feathers, subjects nevertheless accepted the statement "lorches have purple feathers" as true.[26] This is empirical evidence in favor of the theory of stereotypes defended above, because it suggests that subjects have different kinds of generics in mind in the two experiments: when subjects are presented with "lorches have purple feathers" they interpret it as an absolute generic (implying a high proportion of lorches have purple feathers), but when subjects are presented with evidence that only a small proportion of lorches have purple feathers, they nevertheless accept the absolute generic as true. In terms of my analysis, the experiment shows slippage on the value of x.[27]

Essentialist fallacies

A widely held view is that stereotypes involve unwarranted essentializing. This error can be articulated in terms of the natural probability theory of stereotypes. The charge of essentialism is that stereotypes involve beliefs that the feature ascribed to a group in a stereotype is a deep dispositional property of the group, when it in fact is not. This can involve an inference that the frequency with which a group has a feature is a propensity of the group.

---

[26] Cimpian et al. (2010). This is similar to what McKeever and Sterken (2021) call "amplification" of generics. Begby (2013) argues that prejudices are stereotypes which are "epistemically insidious" — once they are acquired, they function as background beliefs which influence how one interprets new evidence about the group in question, perhaps via confirmation bias. Nguyen (2020) similarly notes that the ambiguities associated with generic quantification can facilitate the promulgation of stereotypes. See also Cappelen and Dever (2019).

[27] In another widely-cited study, researchers found that some generics are accepted on the basis of little evidence but "once accepted psychologically they appear to be commonly taken in a rather strong sense." (Abelson and Kanouse 1966, cited in Sterken 2015b and Leslie 2017).

We saw above that stereotypes have a naturalness requirement: a condition of acceptance of a stereotype is that there are worldly constraints such that the group has the property, to some degree at least, non-contingently. Recall this example from above:

(1)   "Russians vote for a person whose first name is Vladimir"
(2)   "Russians are good ballet dancers"
(3)   "Russians have pale skin"

"Russians are good ballet dancers" might wrongly be taken to express an intrinsic and deep fact about Russians, rather than a contingent feature of the historical and cultural milieu of Russia. In general, for many stereotypes such as (2) the facts undergirding them are more like (1) than (3), but many stereotypes are interpreted more like (3) than (1).

So, there can be slippage from contingent facts about a group to beliefs that such facts are more deeply constrained than they really are. Similarly, there can be slippage between the intentions of the assertor of a stereotype and the interlocutor's understanding of the assertor's intentions regarding the degree of worldly constraint of the asserted stereotype. You might assert (2) while implicitly thinking its truth is historically contingent like the truth of (1) is, while I might hear you assert (2) and conclude that its truth is constrained in ways more like the truth of (3) is.[28]

Experiments show that people interpret generic claims in essentialist terms, assuming or concluding that the features ascribed in a generic are natural, intrinsic properties of the group. For example, in one experiment, subjects were given facts expressed in different forms about a hypothetical group of people: generic facts ("Zarpies hate ice cream"), specific facts ("This zarpie hates ice cream"), and unlabeled facts ("This hates ice cream"), and then these subjects were given a series of tests designed to elicit the extent to which they thought that the group (zarpies) had the property (hating ice cream). Those subjects who were given generic facts displayed more category essentialism than other subjects.[29]

To accept a stereotype, it should not only satisfy one of the generic forms articulated in §3, but it should also satisfy the naturalness requirement. To accept a stereotype on only the formal grounds can involve the essentializing fallacy. The essentializing fallacy is widely recognized by scholars of stereotypes. The natural probability theory of stereotypes affords a fresh way of articulating the problem. I argued in §3 that the probabilities in generic claims are more than mere contingent frequencies. The essentializing fallacy can be understood as making an inference from facts about a frequency of a trait in a group to a belief about a deeply-grounded disposition of that group. Though many books are paperbacks, books do not have a deep disposition to be paperbacks; though many Russians have voted for Putin, Russians do not have a non-contingent propensity to vote for a person whose first name is Vladimir.

Of course, not all such inferences are fallacies. Indeed, one source of evidence about a propensity of some entity to have some property is the corresponding frequency

---

[28] See Langton et al. (2012) and Lemeire (2021) for compelling discussions of how such slippage affords "defensive shifting" when an asserter of an offensive stereotype is challenged by counterevidence.
[29] Gelman et al. (2010).

with which we observe that entity manifesting that property. Such evidence is, in some contexts, a reliable guide to propensities. Yet, in addition to such evidence, in many cases a well-grounded inference about a propensity must be based on theoretical knowledge of the entity, and knowledge of how the frequency was determined. To estimate the propensity of this coin to land heads, I could toss it one hundred times and take the frequency of heads as a reliable indicator of its propensity to land heads; in this case the frequency evidence is a good guide to the propensity. But to estimate the propensity of coins in my pocket to be nickels, the proportion (frequency) of nickels now in my pocket would be a poor guide.

As we saw earlier, one of Haslanger's examples of a stereotype that is seemingly supported by frequency data, but that she argues we should nevertheless reject, is "blacks are violent." According to statistics about violent crime in the United States, it appears to be the case that today blacks commit more violent crimes than other racial groups. So, although it is not a true absolute generic, it could be taken as a true relative generic. However, argues Haslanger, we ought to reject the stereotype as false because it assumes that blacks have, relative to other groups, a higher disposition to violence. In my terms, to believe or assert this stereotype involves assuming that the higher frequency of violent crimes among blacks compared to other racial groups is a result of deep constraining facts about these groups. That assumption, Haslanger rightly argues, we should reject. We have no background theory — physical, biological, genetic, sociological — that suggests this difference in frequency of violent crimes is a deeply-rooted difference between groups; conversely, we have overwhelming reasons to think that this difference in frequency of violent crimes is a result of historically contingent (and oppressive and unjust) social circumstances.

7. Cognitive Heuristics or Cognitive Culprits?

Stereotypes have been an active subject of study by social psychology. One of the reasons for the focus of psychology on stereotypes is the extreme harm that stereotypes have caused, particularly in the decades immediately prior to the burgeoning of research in social psychology in the mid-twentieth century. The empirical study of stereotypes has itself been politicized. For example, one prominent social psychologist claims that "as scientists concerned with improving the social condition, we must be wary of arguments that can be used to justify the use of stereotypes."[30] Objecting directly to this, Jussim, another prominent social psychologist, claims that this is a political ambition and thus not properly in the domain of science.[31] Indeed, in recent years there has been a vitriolic debate between some social psychologists, who hold that stereotypes are always false and stereotyping is always wrong, and others, who hold that stereotypes are often accurate and stereotyping is often reasonable. The theory of stereotypes articulated here can resolve this debate.

To some extent this debate may be based on conflicting conceptions of stereotypes. Beeghly distinguishes between a "descriptive" view of stereotypes, which main-

---

[30] Stangor (1995).
[31] Jussim (2012).

tains that stereotypes are just cognitive structures for processing complex social facts, and a "normative" view of stereotypes, which maintains that stereotypes are typically wrong and unjustified.[32] One way to understand the debate between social psychologists is that one side holds a descriptive view of stereotypes while the other side holds a normative view. For example, here is an articulation of the descriptive view: "We define a stereotype as a cognitive structure containing the perceiver's knowledge, beliefs, and expectancies about some human social group."[33] Conversely, here is an articulation of the normative view: "stereotypes are 'nouns that cut slices'; they are the cognitive culprits in prejudice and discrimination."[34]

Nice phrase, *cognitive culprits*. However, one argument for the descriptive view is that it maintains the possibility that some stereotypes can be accurate.[35] On the normative view, stereotypes are deemed false by default; but since not all beliefs or claims about groups are false, the normative view has to say that stereotypes are false beliefs about groups, while true beliefs about groups are something else, not stereotypes. But those true beliefs about groups have the look and feel of stereotypes — they are generic claims attributing properties to human social groups — better (that is, more consistent with linguistic usage and psychological facts) just to say that they are stereotypes. The natural probability theory of stereotypes is a descriptive conception, which directs attention to two kinds of factual questions to assess a given stereotype: whether empirical facts warrant one of the generic forms for the stereotype, and whether empirical or theoretical considerations suggest that this generic is grounded in sufficient worldly constraints such that the naturalness requirement is satisfied (these are not the only grounds on which one can criticize stereotypes, of course — see §8).

The above debate is not merely definitional, however. A key issue in the debate is the extent to which stereotypes are in fact conducive to accurate beliefs about individuals and groups. Many social psychologists claim that stereotypes are inaccurate and contribute to bias. A large amount of empirical work has focused on the negative epistemic and practical influences of stereotypes. For example, empirical studies document a "self-fulfilling prophecy" phenomenon in classrooms, in which teachers' prior expectations about the future performance of particular students modulate that future performance. Such findings are often discussed in conjunction with evidence that teachers tend to have lower prior expectations for racial minorities, for women in math, and for boys in reading. Another prominent example of the negative effect of stereotypes is the phenomenon dubbed "stereotype threat," which occurs when members of a particular social group underperform on skill-based tasks as a result of anxiety that their performance might confirm stereotypes about their group.

Some recent empirical work pushes back against this tradition of studying only the negative consequences of stereotypes. This work studies the extent to which some stereotypes appear to accurately track social facts. One psychologist in this camp claims that "The evidence is clear. Based on rigorous criteria, laypeople's beliefs about groups

---

[32] Beeghly (2015).
[33] Mackie et al. (1996).
[34] Fiske and Taylor (1991).
[35] Jussim (2012).

correspond well with what those groups are really like."[36] The primary method of this camp is to probe the beliefs of subjects regarding social groups, and then to compare these beliefs to evidence gathered from sources such as census data, standardized test scores, or crime statistics. Stereotypes often correspond to these social facts, according to this view, which, claim its proponents, vindicates stereotypes from the excessively negative view described above. On this view, stereotypes are seen as aids to explanation, and as tools for achieving cognitive efficiency which can be deployed to make reliable inferences.[37] Yet, the natural probability account of stereotypes defended here entails that those statistical findings are only half the story: beliefs about social groups can satisfy one of the probabilistic criteria of generics but not the naturalness criterion. It would require much more than the social statistics drawn on by this camp to show, using Jussim's phrase, "what those groups are really like."

Part of this debate involves differing views about the permissibility of stereotyping in general. Some social psychologists explicitly claim that stereotyping is wrong because it involves treating individual people in a non-individualized way, as an identity-less member of a group.[38] Sometimes this position is defended by appealing to normative principles, such as "people should always be treated as individuals," and sometimes the position is defended on epistemic grounds, by claiming that because individuals are all unique, making inferences about individual people by appealing to stereotypes is unreliable.[39] On the other hand, other social psychologists claim that stereotypes are just like other cognitive heuristics, mental devices to help us make sense of a complicated world, and since stereotypes often track social facts, using them to make inferences about individuals can enhance the reliability of those inferences.[40] The central criterion for this latter view is the empirical adequacy of stereotypes, where, as we saw above, that is judged with respect to actual statistical facts about groups.[41]

Thus, there are two competing views about stereotypes among social psychologists. One view holds that stereotypes are always wrong, that stereotypes contribute to bias and injustice, and that stereotypical inferences are bad. The other view holds that stereotypes are very often accurate, stereotypes can contribute to reliable inferences about individuals, and that stereotypical inferences are often fine.

The natural probability theory of stereotypes can help to resolve this dispute. On the one hand, the natural probability theory of stereotypes decisively supports the descriptive view of stereotypes: some stereotypes are false, some are true, and the

---

[36] Jussim (2016); see also Jussim et al. (2009).

[37] McGarty et al. (2004).

[38] Here, for instance, is how John Bargh who studies the psychology of stereotype puts it: stereotyping involves "applying a generalization about a group to an individual, which is always incorrect." Another well-known psychologist of stereotyping (especially well-known for her work on implicit bias), Mahzarin Banaji, claimed that "people should be judged as individuals and not as members of a group" (see Paul 1998).

[39] Of course, it is not just social psychologists who claim that people should always be considered as individuals and that stereotyping fails to do this. For instance, Blum (2004) argues that "a bad of all stereotyping" is the failure to see an individual as an individual. Beeghly (2018) offers a criticism of this position.

[40] See, for example, Arkes and Tetlock (2004).

[41] Jussim et al. (2009).

natural probability theory of stereotypes offers a set of conditions to determine which asserted stereotypes are epistemically warranted and which should be rejected on epistemic grounds (these are necessary conditions for acceptance, not sufficient conditions, because stereotypes can be wrong for non-epistemic, ethical reasons). On the other hand, the natural probability theory of stereotypes supports those who argue that stereotypes that are warranted only by statistical evidence arising from contingent circumstances should be rejected. Such stereotypes are not true in the generic sense but are only true in the statistical sense — they do not satisfy the naturalness requirement for stereotypes — and for that reason should be rejected.

8. Conclusion

I have argued that stereotypes are expressed as generics, and generics can be represented as one of several kinds of statements involving conditional probabilities. One might think that any kind of representational device that can represent generalizations can serve to represent stereotypes. Yet, I noted several properties of generics that make them especially apt for representing stereotypes: that generics are context-sensitive indexicals, and that generic statements imply a type of naturalness. Since these are both central features of stereotypes, so I have argued, generics are especially apt for representing stereotypes. Moreover, some ways of representing generalizations do not seem apt for stereotypes, like statements of mere frequency. To use an example from §2, we should reject generics like "books are paperbacks" as false, though it is true in a merely-statistical sense, and we saw that this is the same reason that Haslanger argues we should reject stereotypes such as "blacks are violent."

I have articulated two ways that asserted stereotypes can fail to be true: by not satisfying one of the generic probabilistic forms, and by not satisfying the naturalness requirement. But this is far from a complete list of the ways that believing or asserting stereotypes can be wrong. For example, Basu argues that a supposedly rational racist — someone who holds racist beliefs which are apparently supported by evidence — nevertheless commits a wrong by harming others.[42] As some argue, moral concerns can encroach on justified belief, and thus even if a stereotype were justified purely on epistemic grounds, moral encroachment could entail that the stereotype is nonetheless unwarranted.[43] Silva gives a different kind of argument, based on Bayesianism, for why sustained belief in some pernicious stereotypes is epistemically unjustified.[44] And Haslanger has argued that asserting stereotypes, even if they are epistemically justified, can generate looping effects which create or exacerbate the conditions under which the stereotype becomes true. Thus, there are many ways that believing or asserting stereotypes can be wrong; the natural probability theory of stereotypes articulates two epistemic ways a stereotype can be wrong.

Beeghly claims that "we have little reason to build moral or epistemic defect into the very idea of a stereotype."[45] While it may be true that stereotypes ought not be stip-

---

[42] Basu (2019).
[43] See, for example, Gardiner (2018).
[44] Silva (2020).
[45] Beeghly (2015).

ulated as epistemically problematic, I have argued that there are two kinds of epistemic errors routinely associated with stereotypes. Plausibly many stereotypes have one or both of these epistemic defects. In addition to various moral grounds for sanctioning stereotypes, the natural probability theory of stereotypes articulates two kinds of possible epistemic defects of stereotypes that can be appealed to for their sanction.

Recent work on the semantics of generics and the psychology of reasoning about categories lends support to what I am calling the natural probability theory of stereotypes. Stereotypes are expressed as generics, taking one of several forms involving conditional probabilities, in which the probabilities are understood as dispositions resulting from constraining facts about the group in question. This theory of stereotypes predicts two families of fallacies associated with stereotypes and stereotypical reasoning: one based on fallacious probabilistic reasoning, and one based on fallacious essentialist reasoning. Empirical findings in psychology suggest that people often commit these fallacies, which could in part explain the tenacity of epistemically unwarranted stereotypes. The natural probability theory of stereotypes also helps to resolve an active debate in social psychology between those researchers who claim that stereotypes are always wrong and stereotyping is always bad, and other researchers who claim that stereotypes are just cognitive heuristics, to be judged only on their empirical merits.

Conflict of Interest: The author declares no conflict of interest.

References

Abelson R., Kanouse D. (1966), "Subjective Acceptance of Verbal Generalizations," [in:] *Cognitive Consistency: Motivational Antecedents and Behavioral Consequents*, S. Feldman (ed.), Academic Press, New York, pp. 171–197.

Arkes H., Tetlock P. (2004), "Attributions of Implicit Prejudice, or 'Would Jesse Jackson 'Fail' the Implicit Association Test?'," *Psychological Inquiry* 15 (4): 257–278.

Asher N., Pelletier J. (2012), "More Truths About Generic Truth," [in:] *Genericity*, A. Mari, C. Beyssade, F.D. Prete (eds.), Oxford University Press, Oxford: 313–333.

Basu R. (2019), "The Wrongs of Racist Beliefs," *Philosophical Studies* 176 (9): 2497–2515.

Beeghly E. (2015), "What is a Stereotype? What is Stereotyping?," *Hypatia* 30 (4): 675–691.

Beeghly E. (2018), "Failing to Treat Persons as Individuals," *Ergo* 5 (26): 687–711.

Begby E. (2013), "The Epistemology of Prejudice," *Thought: A Journal of Philosophy* 2 (2): 90–99.

Berkovitz J. (2015), "The Propensity Interpretation of Probability: A Re-evaluation," *Erkenntnis* 80: 629–711.

Blum L. (2004), "Stereotypes and Stereotyping: A Moral Analysis," *Philosophical Papers* 33 (3): 251–289.

Bosse A. (2022), "Stereotyping and Generics," *Inquiry: An Interdisciplinary Journal of Philosophy*: 1–17.

Brandone A.C., Gelman S.A., Hedglen J. (2015), "Children's Developing Intuitions About the Truth Conditions and Implications of Novel Generics Versus Quantified Statements," *Cognitive Science* 39 (4): 711–738.

Cappelen H., Dever J. (2019), *Bad Language*, Oxford University Press, Oxford.

Chestnut E., Markman E. (2018), "'Girls Are as Good as Boys at Math' Implies That Boys Are Probably Better: A Study of Expressions of Gender Equality," *Cognitive Science* 42 (7): 2229–2249.

Cimpian A., Brandone A.C., Gelman S.A. (2010), "Generic Statements Require Little Evidence for Acceptance but have Powerful Implications," *Cognitive Science* 34 (8): 1452–1482.

Cohen A. (1999), "Generics, Frequency Adverbs, and Probability," *Linguistics and Philosophy* 22 (3): 221–253.

Fiske S.T., Taylor S.E. (1991), *Social Cognition*, McGraw-Hill, New York.

Foster-Hanson E., Leslie S.J., Rhodes M. (2016), "How Does Generic Language Elicit Essentialist Beliefs?," [in:] *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, A. Papafragou, D. Grodner, D. Mirman, J.C. Trueswell (eds.), Cognitive Science Society, Philadelphia, PA, pp. 1541–1546.

Gardiner G. (2018), "Evidentialism and Moral Encroachment," in *Believing in Accordance with the Evidence*, K. McCain (ed.), Springer Verlag.

Gelman S.A., Ware E.A., Kleinberg F. (2010), "Effects of Generic Language on Category Content and Structure," *Cognitive Psychology* 61 (3): 273–301.

Hammond M.D., Cimpian A. (2017), "Investigating the Cognitive Structure of Stereotypes: Generic Beliefs About Groups Predict Social Judgments Better Than Statistical Beliefs," *Journal of Experimental Psychology: General* 146 (5): 607–614.

Haslanger S. (2012), *Resisting Reality: Social Construction and Social Critique*, Oxford University Press, Oxford.

Humphreys P. (1985), "Why Propensities Cannot Be Probabilities," *Philosophical Review* 94 (4): 557–570.

Jussim L. (2012), *Social Perception and Social Reality: Why accuracy dominates bias and self-fulfilling prophecy*, Oxford University Press, Oxford.

Jussim L. (2016), "Truth in Stereotypes," *Aeon*: August 15.

Jussim L., Cain T.R., Crawford J.T., Harber K., Cohen F. (2009), "The Unbearable Accuracy of Stereotypes," [in:] *Handbook of Prejudice, Stereotyping, and Discrimination*, T.D. Nelson (ed.), Psychology Press, New York: 199–227.

Langton R., Haslanger S., Anderson L. (2012), "Language and Race," [in:] *Routledge Companion to Philosophy of Language*, G. Russell, D. Graff Fara (eds.), Routledge, London: 753–767.

Lemeire O. (2021), "Falsifying Generic Stereotypes," *Philosophical Studies* 178 (7): 2293–2312.

Leslie S.-J. (2008), "Generics: Cognition and Acquisition," *Philosophical Review* 117 (1): 1–47.

Leslie S.-J. (2016), "Generic Generalizations," [in:] *The Stanford Encyclopedia of Philosophy*, E.N. Zalta (ed.), URL = https://plato.stanford.edu/entries/generics/ [Accessed 03.09.2023].

Leslie S.-J. (2017), "The Original Sin of Cognition: Fear Prejudice, and Generalisation," *The Journal of Philosophy* 114 (8): 393–421.

Mackie D.M., Hamilton D.L., Susskind J., Rosselli F. (1996), "Social Psychological Foundations of Stereotype Formation," [in:] *Stereotypes and Stereotyping*, C.N. Macrae, C. Stengor, M. Hewstone (eds.), Guilford Press, New York: pp. 41–78.

McGarty C., Yzerbyt V., Spears R. (2004), *Stereotypes as Explanations: The Formation of Meaningful Beliefs About Social Groups*, Cambridge University Press, Cambridge.

McKeever M., Sterken R. (2021), "Social and Political Aspects of Generic Language and Speech," [in:] *The Routledge Handbook of Social and Political Philosophy of Language*, J. Khoo, R. Sterken (eds.), Routledge, pp. 259–280.

Nguyen A. (2020), "The Radical Account of Bare Plural Generics," *Philosophical Studies* 177 (5): 1303–1331.

Nickel B. (2012), "Saying and Doing: The Role of Semantics in the Use of Generic Sentences," *The Canadian Journal of Linguistics* 57 (2): 289–302.

Nickel B. (2016), *Between Logic and the World: An Integrated Theory of Generics*, Oxford University Press, Oxford.

Paul A.M. (1998), "Where Bias Begins: The Truth About Stereotypes," *Psychology Today* 31 (3): 52.

Pelletier F.J., Asher N. (1997), "Generics and Defaults," [in:] *Handbook of Logic and Language*, J. van Benthem, A. Ter Meulen (eds.), Elsevier/The MIT Press, Amsterdam, Cambridge, pp. 1125–1177.

Rhodes M., Leslie S.J., Tworek C. (2012), "The Cultural Transmission of Social Essentialism," *Proceedings of the National Academy of Sciences* 109 (34): 13526–13531.

Rosch E. (1973), "Natural Categories," *Cognitive Psychology* 4 (3): 328–350.

Silva P. (2020), "A Bayesian Explanation of the Irrationality of Sexist and Racist Beliefs Involving Generic Content," *Synthese* 197 (6): 2465–2487.

Slater M. (2015), "Natural Kindness," *British Journal for the Philosophy of Science* 66 (2): 375–411.

Stangor C. (1995), "Content and Application Inaccuracy in Social Stereotyping," [in:] *Stereotype Accuracy*, Y.-T. Lee, L.J. Jussim, C.R. McCauley (eds.), American Psychological Association, Washington DC, pp. 275–292.

Sterken R.K. (2015a), "Generics in Context," *Philosophers' Imprint* 15 (21): 1–30.

Sterken R.K. (2015b), "Generics, Content and Cognitive Bias," *Analytic Philosophy* 56 (1): 75–93.

Sterken R.K. (2017), "The Meaning of Generics," *Philosophy Compass* 12 (8): e12431.

Suarez M. (2020), *Philosophy of Probability and Statistical Modelling*, Cambridge Elements in Philosophy of Science, Cambridge University Press, Cambridge.

Tessler M.H., Goodman N.D. (2019), "The Language of Generalization," *Psychological Review* 126 (3): 395–436.

van Rooij R. (2019), "Generics and Typicality," *Proceedings of Sinn Und Bedeutung* 22 (2): 443–461.